

SUSMAN GODFREY L.L.P.

A REGISTERED LIMITED LIABILITY PARTNERSHIP
ONE MANHATTAN WEST
NEW YORK, NEW YORK 10001-8602
(212) 336-8330
FAX (212) 336-8340
WWW.SUSMANGODFREY.COM

SUITE 5100
1000 LOUISIANA STREET
HOUSTON, TX 77002-5096
(713) 651-9366

SUITE 1400
1900 AVENUE OF THE AMERICAS
LOS ANGELES, CALIFORNIA 90067-6029
(310) 789-3100

SUITE 3000
401 UNION STREET
SEATTLE, WASHINGTON 98101-2683
(206) 516-3880

IAN B. CROSBY
DIRECT DIAL (206) 516-3861

E-MAIL ICROSBY@SUSMANGODFREY.COM

November 6, 2024

Hon. Ona T. Wang
Daniel Patrick Moynihan
United States Courthouse
500 Pearl Street
New York, NY 10007-1312

Re: *The New York Times Co. v. Microsoft Corp., et al.* No. 1:23-cv-11195
New York Daily News, LP, et al. v. Microsoft Corp., et al., No. 1:24-cv-3285
The Center for Investigative Reporting, Inc. v. OpenAI, Inc., No. 1:24-cv-04872

Dear Magistrate Judge Wang:

Per the Court's request during the October 30th status conference, on behalf of all parties in the above-referenced actions, counsel for The Times hereby submits the below chart containing a summary of discovery disputes for discussion at the upcoming December 3rd conference.

As the Court instructed, the parties will submit an updated version of this chart by November 22nd outlining any disputes that have ripened since this initial filing.

Sincerely,

/s/ Ian B. Crosby

Ian B. Crosby

Key: Green represents Plaintiffs' filing; Purple represents Defendants' filing; Orange represents joint filing

Motion Dkt. No.	Opposition Dkt. No.	Topic in Dispute	Party Raising Dispute & Date Motion Filed	Status of Dispute	Parties' Shorthand Positions
Disputes Related to OpenAI's Alleged Use of Publisher Content					
283 ¹	301	The Times's RFPs 84–85 (documents concerning OpenAI's knowledge and use of Dnyuz and similar websites that copy publisher content without permission)	The New York Times, October 24, 2024	NYT and OpenAI have submitted letter briefs, which the Court has not yet heard argument on.	<p>Plaintiff's position: The dispute as to RFP 84 has been resolved. OpenAI agrees to “review and produce responsive, non-privileged documents” in response to this RFP, which seeks documents concerning OpenAI's knowledge of and use of Dnyuz, a known plagiarizer of publisher content. <i>See</i> Dkt. 301 at 1. As to RFP 85, The Times requests an order compelling OpenAI to produce documents in response to RFP 85, which seeks documents concerning OpenAI's knowledge and use of websites similar to Dnyuz that likewise copy publisher content without permission. This Request is relevant to The Times's allegations of willful infringement, among other issues. FAC ¶¶ 124-26.</p> <p>OpenAI's position: There is no dispute regarding RFP 84 because OpenAI has already run “Dnyuz” as a search term and has been reviewing and producing responsive documents. Regarding RFP 85, OpenAI objected to “websites besides Dnyuz” as overbroad and vague. OpenAI is willing to search for websites The Times considers similar to Dnyuz if The Times provides a list of such websites.</p>
283	301	The Times's RFP 26 (documents concerning OpenAI's use of web crawlers)	The New York Times, October 24, 2024	NYT and OpenAI have submitted letter briefs, which the Court has not yet heard argument on.	<p>Plaintiff's position: The Times requests an order compelling OpenAI to produce documents concerning web crawlers used by Microsoft or by developers of Custom GPTs. The News Plaintiffs allege that OpenAI uses such crawlers to, without permission, gather copies of publisher works for use in Defendants' generative AI products. <i>See</i> Daily News Compl. ¶ 147; <i>see also</i> Ex. H (Dkt. 283-8) (agreement produced in discovery that supports these allegations).</p> <p>OpenAI's position: OpenAI already agreed to conduct a reasonable search for and produce documents concerning OpenAI's use of web crawlers to crawl content of articles published by The Times in connection with training the relevant models. OpenAI has no control</p>

¹ Unless otherwise indicated, all docket numbers are from *New York Times Company v. Microsoft Corp., et al.*, Case No. 1:23-cv-11195-SHS.

Motion Dkt. No.	Opposition Dkt. No.	Topic in Dispute	Party Raising Dispute & Date Motion Filed	Status of Dispute	Parties' Shorthand Positions
					over how Microsoft or third-party GPT developers use web crawlers. Nonetheless, OpenAI remains committed to meeting and conferring about this RFP.
283	301	The Times's RFP 18 (documents concerning removal of copyright management information)	The New York Times, October 24, 2024	NYT and OpenAI have submitted letter briefs, which the Court has not yet heard argument on.	<p>Plaintiff's position: The Times requests an order compelling OpenAI to produce documents concerning any removal of copyright management information (CMI) in connection with retrieval-augmented generation (RAG). OpenAI has agreed to produce documents concerning the removal of CMI in connection with model training but has refused to produce documents concerning the removal of CMI in connection with RAG, a technology that enables OpenAI to integrate live Times content into its generative AI products. This request is relevant to The Times's DMCA claim, which includes claims for removing or altering CMI, including in connection with RAG. <i>See</i> FAC ¶¶ 181-91.</p> <p>OpenAI's position: Documents the Times is now requesting are not requested in RFP 18, which seeks documents only about alleged CMI removal from training datasets, not in connection with RAG. The Times should serve an RFP directed to CMI removal in connection with RAG so that OpenAI may properly consider, object, and respond to it.</p>
Disputes Related to Plaintiffs' Requests for Inspection					
250	250	OpenAI's Model Inspection Protocol	News Plaintiffs, September 27, 2024	Parties outlined their positions in September 27 status letter and await the Court's ruling	<p>Plaintiff's position: The Times and Daily News Plaintiffs request an order compelling OpenAI to provide Plaintiffs sufficient retail "credits" to analyze the extent to which unauthorized copies of Plaintiffs' copyrighted works are encoded within the parameters of the GPT LLMs themselves and the propensity of the GPT LLMs to disseminate or publicly display those works in outputs. The News Plaintiffs further request that the Court reject OpenAI's proposal to profit from Plaintiffs' discovery by charging Plaintiffs what it charges its customers for retail services as part of its (for profit) business.</p> <p>OpenAI's position: OpenAI has already agreed to provide Plaintiffs \$15,000 each in API credits, and to negotiate in good faith about any</p>

Motion Dkt. No.	Opposition Dkt. No.	Topic in Dispute	Party Raising Dispute & Date Motion Filed	Status of Dispute	Parties' Shorthand Positions
					future credits that Plaintiffs may need with the expectation that the parties split the cost evenly. This iterative approach will provide the parties real world experience with running efficient searches, and also ensures that Plaintiffs are not paying "retail" pricing. Plaintiffs have not conducted any model inspections to date and have not used any of the \$15,000 API credits provided. Thus, their demand for \$800,000 in credits now is premature.
305	305	OpenAI's Training Data	News Plaintiffs, November 1, 2024	Submitted joint dispute letter per Court's order (Dkt. 304)	<p>Plaintiff's position: The Times and the Daily News Plaintiffs request an order compelling OpenAI to identify and admit which of Plaintiffs' works were used to train each of the GPT-1, GPT-2, GPT-3, GPT-3.5, GPT-3.5 Turbo, GPT-4, GPT-4 Turbo, and GPT-4o families of models. Because this information is solely within Defendants' possession, News Plaintiffs have served numerous discovery requests since February seeking this information, and now request that OpenAI be ordered to provide this information instead of requiring Plaintiffs and their experts to spend weeks upon weeks in their tightly controlled "sandbox."</p> <p>OpenAI's position: Plaintiffs recently served an effective total of almost 500 million requests for admission, for which OpenAI's responses are not even due yet. The information that Plaintiffs seek is not solely in OpenAI's hands because OpenAI made training data available to Plaintiffs in early June (though Plaintiffs did not schedule their first inspection until October). There is no "sandbox." Rather, to accommodate Plaintiffs' requests to access several hundred terabytes of unstructured textual training data, OpenAI built an enterprise-grade virtual machine and installed hundreds of software tools for Plaintiffs' exclusive use. OpenAI has resolved Plaintiffs' concerns about their technical inspection of training data and continues to meet and confer about providing more technical assistance and support.</p>
Disputes Related to the Scope of Products at Issue					
279	294	Scope of Microsoft's products subject to discovery	News Plaintiffs, October 23, 2024	NYT and Microsoft have submitted letter briefs, which the Court has not yet	Plaintiff's position: The News Plaintiffs request an order compelling Microsoft to search for and produce 1) documents responsive to The Times's RFP 67 regarding its plans to commercialize existing generative AI products built with Times content, including but not limited to Bing Chat and Microsoft

Motion Dkt. No.	Opposition Dkt. No.	Topic in Dispute	Party Raising Dispute & Date Motion Filed	Status of Dispute	Parties' Shorthand Positions
				heard argument on	<p>Copilot; and 2) documents concerning Copilot Daily—Microsoft's recently announced AI-powered news product—and any other generative AI products that Microsoft is currently developing in categories that compete with News Plaintiffs' products—namely news, cooking, consumer reviews, and sports. News Plaintiffs allege an ongoing and continuous course of illegal conduct, which entitles them to discovery on Microsoft's post-complaint conduct, including with respect to products that remain in development.</p> <p>Microsoft's position: Microsoft requests a protective order that the scope of discovery encompasses only those products released to the public at the time Plaintiffs' respective complaints were filed and actually identified in the complaints. Discovery into products released after the lawsuits were filed, and into new products that may or may not even be released, should be denied. In particular, discovery regarding Copilot Daily should be denied, as that product is based on an entirely new technology stack and business model not named in the complaint, and that was only released a few weeks ago in October.</p>
287	302	Scope of OpenAI's products subject to discovery	News Plaintiffs, October 24, 2024	NYT and OpenAI have submitted letter briefs, which the Court has not yet heard argument on	<p>Plaintiff's position: The News Plaintiffs seek an order compelling OpenAI to search for and produce documents concerning SearchGPT (a GPT-based search product powered by News Plaintiffs' content), the OpenAI API, Copilot, and any other current or anticipated GPT-based product (including CustomGPTs) that relate to content in categories that compete with News Plaintiffs' products—namely news, cooking, consumer reviews, and sports. News Plaintiffs allege an ongoing and continuous course of illegal conduct, which entitles them to discovery on OpenAI's post-complaint conduct, including with respect to anticipated GPT-based products.</p> <p>OpenAI's position: OpenAI has already agreed to produce documents about every specific product Plaintiffs have identified, i.e., ChatGPT (including Browse (f/k/a Browse with Bing)), SearchGPT, the OpenAI API, and Microsoft's Copilot. The scope of Plaintiffs' request for documents concerning "any other current or anticipated GPT-based product (including CustomGPTs)" is overly</p>

Motion Dkt. No.	Opposition Dkt. No.	Topic in Dispute	Party Raising Dispute & Date Motion Filed	Status of Dispute	Parties' Shorthand Positions
					broad and vague, as it encompasses a sweeping, undefined and unknown range of third-party products and other product categories. OpenAI remains committed to continuing to meet and confer regarding Plaintiffs' requests.
Disputes Related to Custodians and Custodial Documents					
204	211	OpenAI's custodians (The Times seeks eight additional OpenAI custodians)	The New York Times, August 12, 2024	Addressed at the October 30th conference; OpenAI ordered to provide hit counts for the eight disputed custodians by November 13 (Dkt. 304)	<p>Plaintiff's position: The News Plaintiffs stand by their request for the following eight additional OpenAI custodians: James Betker, William Fedus, Daniel, Kokotajlo, Brad Lightcap, Ryan Lowe, William Saunders, Ilya Sutskever, and Wojciech Zaremba. The News Plaintiffs will review the hit counts that OpenAI provides by November 13 and confer with OpenAI.</p> <p>OpenAI's position: OpenAI will compile hit counts pursuant to the Court's order. As The Times completes its review of documents from OpenAI's 26 agreed-upon custodians, the parties will continue to meet and confer about any alleged gaps in OpenAI's productions including the need for any additional custodians.</p>
228	233	OpenAI's search terms (The Times requests that OpenAI run additional search terms for its custodians' ESI)	The New York Times, August 27, 2024	Addressed at the October 30th conference; the Court ruled that the total hit count will be capped at or around 500,000 for the current set of agreed custodians (Dkt. 304) and the parties will meet and confer about the precise search terms and rolling productions	<p>Plaintiff's position: The News Plaintiffs are preparing a revised search term proposal intended to capture around 500,000 documents for the current set of agreed custodians.</p> <p>OpenAI's position: The parties are continuing to meet and confer about search terms following the Court's guidance at the October 30 conference setting a total hit count cap of 500,000 documents for the current set of agreed custodians.</p>

Motion Dkt. No.	Opposition Dkt. No.	Topic in Dispute	Party Raising Dispute & Date Motion Filed	Status of Dispute	Parties' Shorthand Positions
Disputes Related to Licensing and Data Access Documents					
141	147	The Times's RFP 6 (communications regarding licensing and data agreements)	The New York Times, June 11, 2024	NYT and OpenAI have submitted letter briefs, which the Court heard argument on at the September 12 conference; parties ordered to further meet and confer, and the parties continue to do so	<p>Plaintiff's position: The Times seeks an order requiring OpenAI to produce internal and external communications relating to agreements in which it licenses content to use with its generative AI models and products. The parties are close to resolving this dispute. The sole remaining issue is whether these productions should be limited to the currently agreed upon 26 custodians, or whether OpenAI should also produce documents from other employees involved in these licensing negotiations. The parties are continuing to meet and confer about this disagreement and hope to resolve it before the December 3 conference.</p> <p>OpenAI's position: OpenAI has produced all relevant data access agreements for which third party confidentiality concerns have been resolved. OpenAI has also agreed to produce non-privileged relevant communications between (a) OpenAI and third-parties related to data access agreement negotiations, and (b) between OpenAI employees related to data access agreement negotiations. OpenAI is committed to continuing to meet and confer about any alleged gaps in OpenAI's productions, including the need for additional custodians, after The Times completes its review of such documents.</p>
Disputes Related to Documents Concerning OpenAI's Alleged Transition to a For-Profit Company					
141	147	The Times's RFP 11 (OpenAI's alleged transition to a for-profit company)	The New York Times, June 11, 2024	NYT and OpenAI have submitted letter briefs, which the Court heard argument on at the September 12 conference; parties ordered to further meet and confer and the parties continue to do so	<p>Plaintiff's position: The Times seeks an order requiring OpenAI to produce documents relating to its transition to a for-profit company. The parties are close to resolving this dispute. The sole remaining issue concerns the extent to which OpenAI must log communications with lawyers. The parties are continuing to meet and confer and hope to resolve this dispute before the December 3 conference.</p> <p>OpenAI's position: OpenAI has agreed to a conduct a reasonable search of its designated custodians and produce documents relating to OpenAI's creation of a for-profit entity (regardless of whether those documents are communications with Microsoft) to the extent those documents address the operation of OpenAI's relevant business or, as a general matter, its ability to raise capital. The Times</p>

Motion Dkt. No.	Opposition Dkt. No.	Topic in Dispute	Party Raising Dispute & Date Motion Filed	Status of Dispute	Parties' Shorthand Positions
					had proposed, and OpenAI agreed, that OpenAI need not review or log any communications between OpenAI's outside counsel and in-house counsel that are solely responsive to RFP 11.
Disputes Related to Depositions					
306	306	Early 30(b)(6) custodial depositions	All parties, November 1, 2024	Submitted joint dispute letter per Court's order (Dkt. 304)	<p>Plaintiff's position: The News Plaintiffs seek two early 30(b)(6) custodial depositions prior to the upcoming December 3, 2024 discovery conference. Specifically, Plaintiffs seek one deposition of OpenAI and one deposition of Microsoft, each limited to 7 hours, and each shared with the Class Plaintiffs in the <i>Authors Guild</i> case. These depositions will "help focus document discovery" and resolve many of the parties' disputes. <i>See</i> 10/30/24 Tr. 8:2-20.</p> <p>Microsoft's position: There is no basis for a custodial deposition of Microsoft. Plaintiffs have raised no issue with the court relating to Microsoft's custodians, sufficiency of document collection, or document retention, so there is no basis for discovery about discovery from Microsoft. To the extent Plaintiffs seek a deposition relating to output data, that is not a "custodial" deposition; it is a merits deposition. And Plaintiffs did not previously request it; they have no grounds for taking it on short notice before the next conference; and such an accelerated schedule would pose a substantial burden. Additionally, this issue was never fully briefed as Microsoft had $\frac{3}{4}$ of a page to address its position in the joint letter filed with the Court. Nor did Plaintiffs serve any notice, nor seek to meet and confer about any of their proposed deposition topics before the Status Conference. The Court should deny Plaintiffs' requested relief as to Microsoft.</p> <p>OpenAI's position: OpenAI is conferring with Plaintiffs regarding early custodial 30(b)(6) depositions, but Plaintiffs are seeking a deposition from OpenAI on overbroad, privileged and improper topics, while refusing to provide any custodial 30(b)(6) themselves. The parties' early custodial 30(b)(6) depositions should be mutual, cabined to 7 truly custodial topics per side, limited to 3-hours each, and take place in January 2025.</p>

Motion Dkt. No.	Opposition Dkt. No.	Topic in Dispute	Party Raising Dispute & Date Motion Filed	Status of Dispute	Parties' Shorthand Positions
Disputes Related to Plaintiffs' Search for and Production of Fair Use Documents					
276	293	OAI's RFPs 35 & 37 (Times' analyses of its trends and performance)	OpenAI, October 22, 2024	OpenAI and NYT have submitted letter briefs. The Court has not yet heard full argument and instructed the parties to address the issue in this chart.	<p>OpenAI's position: OpenAI seeks production of (1) documents and communications relating to formal or informal analyses of trends in readership or online subscriptions, in addition to the causes thereof; and (2) documents and communications relating to formal or informal analyses of the performance of content published on the Times' website, including analyses of methods to evaluate that performance. These materials relate directly to the harm alleged in the complaint, and they are important for OpenAI's fair use defense.</p> <p>Plaintiff's position: The Times agreed to produce (1) documents sufficient to show trends in readership and online subscriptions (RFP 35) and (2) documents sufficient to show trends in revenue from The Times's online content (RFP 37). Separately, The Times agreed to produce analyses regarding whether decreases in traffic to The Times's website were caused by Defendants' generative AI products or other market forces, as well as all documents and communications regarding the harm alleged in this case. This discovery is more than sufficient for OpenAI to assess The Times's damages allegations and the effect of OpenAI's infringement on the potential market for The Times's copyrighted works, which is the fourth fair use factor.</p>
276	293	OAI RFP 70 (Benefits of OpenAI's technology)	OAI, October 22, 2024	OpenAI and NYT have submitted letter briefs.	<p>OpenAI's position: OpenAI seeks production of documents and communications concerning the benefits of OpenAI's technology to the journalism industry. These materials relate to OpenAI's fair use defense and demonstrate the public benefits of OpenAI's technology.</p> <p>Plaintiff's position: The "public benefits" inquiry of the fourth fair use factor looks at the public benefits of Defendants' use of Times works. The Times already agreed to produce documents concerning Defendants' use of Times works, which will include documents about any purported "benefits" of Defendants' use, if any exist. Anything else OpenAI seeks in this discovery request is irrelevant and overbroad.</p>

Motion Dkt. No.	Opposition Dkt. No.	Topic in Dispute	Party Raising Dispute & Date Motion Filed	Status of Dispute	Parties' Shorthand Positions
276	293	OAI RFP 76 & 82 (impact of AI and use of AI on the Times)	OAI, October 22, 2024	OpenAI and NYT have submitted full briefing.	<p>OpenAI's position: OpenAI seeks production of (1) documents regarding the impact of third-party generative AI products on journalism (that is, products offered by companies other than OpenAI), and (2) the Times' own use of AI products in support of its business. This topic overlaps to some extent with OpenAI's prior motion to compel (Docket No. 263), which is under submission. As with the issues raised in that prior motion, the requested documents are relevant to OpenAI's fair use defense.</p> <p>Plaintiff's position: Discovery into (1) the impact of third-party generative AI tools on the journalism market and (2) The Times's use of third-party generative AI tools for marketing is not relevant to OpenAI's fair use defense. The fair use inquiry is context-specific and is centered on <i>Defendants'</i> use of The Times's proprietary works and any public benefits that use may produce, a topic on which The Times has already agreed to produce documents.</p>
Protective Order Matters					
<i>Daily News v. Microsoft Corp.</i> , Case No. 1:24-cv-3285-SHS-OTW, Dkt. 163	<i>Daily News v. Microsoft Corp.</i> , Case No. 1:24-cv-3285-SHS-OTW, Dkt. 172	OpenAI's Opposition to the News Plaintiffs' Disclosure of Highly Confidential Material to Dr. Baeza-Yates	OpenAI, October 16, 2024	OpenAI and NYT submitted their positions in the October 16 Status Conference agenda ahead of the October 31 conference	<p>OpenAI's position: OpenAI objects to the disclosure of its highly confidential, internal technical and product-related documents (including source code) to Plaintiffs' expert, Dr. Baeza-Yates, given his role as the co-founder and Chief Scientific Officer for a commercial enterprise—Theodora.AI—that, like OpenAI, is engaged in the business of “fine-tuning” large language models. OpenAI seeks protection from disclosure given the highly commercially sensitive nature of the OpenAI internal documents and source code at issue.</p> <p>Plaintiff's position: OpenAI objects to the disclosure of its highly confidential information to Dr. Baeza-Yates because he is a co-founder and officer of Theodora.AI, an entity “engaged in the business of fine-tuning large language models.” In a case where large language models are at issue, it is unreasonable for OpenAI to oppose disclosing its highly confidential information to Plaintiffs' technical expert simply because he works in the field. OpenAI has not argued that Theodora.AI—a six-person Chilean company primarily focused on detecting bias in Spanish Language text—is a competitor. Thus, OpenAI has not demonstrated that disclosing its</p>

Motion Dkt. No.	Opposition Dkt. No.	Topic in Dispute	Party Raising Dispute & Date Motion Filed	Status of Dispute	Parties' Shorthand Positions
					highly confidential information to Dr. Baeza-Yates presents any risk of competitive harm.